doi:10.11959/j.issn.1000-436x.2015325

机会网络中用户属性隐私安全的高效协作者资料匹配协议

李永凯[1,2]，刘树波[1,2]，杨召唤[1,2]，刘梦君[1,2]
(1. 武汉大学 计算机学院，湖北 武汉 430072；2. 空天信息安全与可信计算教育部重点实验室，湖北 武汉 430072)

摘 要：在机会网络中，用户通过移动造成的相遇性机会，借助协作者实现消息的传输与内容的共享。为了克服现有协作者匹配协议加解密效率不高的问题，针对机会网络中用户的不同隐私要求，设计了 3 个不依赖同态加密的高效隐私内积计算协议，可以证明所提出的协议是隐私安全并且正确的。在此基础上，对所提出的 3 个协议的计算开销与通信开销，与现有工作进行了理论上的比较。仿真结果表明，所提协议能够高效地完成隐私安全匹配，其加解密时间要比基于 Paillier 加密体系的协议要少至少一个数量级。

关键词：机会网络；协作者匹配；隐私安全协议；资料匹配
中图分类号：TP309.2 **文献标识码**：A

Efficient and privacy-preserving profile matching protocols in opportunistic networks

LI Yong-kai[1,2], LIU Shu-bo[1,2],YANG Zhao-huan[1,2],LIU Meng-jun[1,2]
(1.School of Computer, Wuhan University, Wuhan 430072, China;
2.Key Laboratory of Aerospace Information Security and Trusted Computing, Ministry of Education, Wuhan 430072, China)

Abstract: In opportunistic networks, users can take the advantage of parallel opportunistic contacts with other mobile users and find the appropriate helpers to forward the messages or share the contents. Most of the existing profile-matching protocols were designed on the basis of homomorphic cryptosystem and were not quite efficient in encryption and decryption. Three efficient and privacy-preserving profile matching protocols were proposed to deal with different privacy requirements in opportunistic networks, which do not use any homomorphic encryption. The proposed protocols were proved to be privacy-preserving and correct. The performances of the protocols are thoroughly analyzed and evaluated via real smartphone experiments, and the results show that the proposed protocols can decrease encryption and decryption time by at least an order of magnitude than the Paillier cryptosystem based protocol.

Key words: opportunistic network; help-node finding; privacy-preserving protocol; profile matching

1 引言

随着大量具备短距离无线通信能力的智能设备的出现与发展，移动自组网（MANET, mobile ad hoc network）得到了越来越多的关注[1~3]。机会网络作为 MANET 中极具应用价值的方向之一[4]，已成为近年来的研究热点,并被广泛应用于移动社交、移动医疗、参与感知以及位置服务等新兴领域[5~8]。

在机会网络中，移动设备利用彼此之间的“偶遇”产生的通信机会来交换信息、传递文件或者共享服务[9,10]。当前，在该领域中大多数研究都以所有节点协作为前提条件。而在实际应用中，由于移动节点普遍属于不同实体，出于隐私性考虑亟需研究一般的协作建立机制以保证机会计算的服务质量。一个比较合理的解决方案是用户通过相互间兴趣、病症、关键词、可用服务等的匹配建立特定协作。这里主要有 2 种基本类型的匹配：1）基于请求用户与被请求用户之间可用服务、可用应用等的

收稿日期：2015-03-25；**修回日期**：2015-11-04
基金项目：国家重点基础研究发展计划（“973”计划）基金资助项目（2011CB302306）；国家自然科学基金资助项目（41371402）
Foundation Items: The National Basic Research Program of China(973 Program) (2011CB302306); The National Natural Science Foundation of China (41371402)

匹配，这里只有被请求用户满足请求用户的特定需求时，用户间才能建立相应协作，例如，在参与感知服务体系中，发起者与拥有与之匹配的传感器应用的用户建立协作，借助协作者完成传感任务[11]；2）基于请求用户与被请求用户之间兴趣、病症等的相似程度的匹配，这里当用户间的相似程度大于设定的阈值时，用户间建立特定协作，例如，在移动社交网络中，发起者希望寻求具有一致兴趣的协作用户来共享信息[12]。由于用户的兴趣、病症、可用服务等涉及到用户的隐私信息，如何隐秘而又安全地通过匹配来发现自己周围合适的协作用户已经为科研工作者关注的焦点[12,13]。

明显地，第一种类型的协作者匹配的隐私性比第二种类型的隐私性要求低。现阶段关于协作者匹配的研究大多把隐秘匹配问题转化为隐秘向量内积的计算问题。对于第一种类型的协作者隐秘匹配，这里要求在匹配协议结束时发起者了解到其与协作者的共同属性，这种匹配类似于隐秘集合交集的计算。而隐私集合交集大多基于集合多项式和同态加密的思想，其计算开销较大[14]。同时，Dong 等[15]指出隐秘集合交集的计算开销不会小于隐私内积的计算开销。因此如何利用隐私内积计算设计一个高效的并且能够满足第一种匹配需求的隐私安全匹配协议是现阶段所面临的一个挑战。

由于第二种类型协作者匹配的应用涉及移动社交网络、移动医疗以及机会网络中的数据传输等领域，因此现阶段的成果较多[7,12,13,16~18]。例如，针对移动医疗社交网络这个特殊的应用场景，Lu 等[17]提出的 SPOC 框架中以病人所共同拥有的疾病数目作为匹配条件来寻找协作者，从而用户可以借助协作者的智能终端应对紧急情况。另外，许多文献提出了基于秘密共享以及同态加密的隐私安全用户资料（兴趣、病症等）匹配方法[12,13,16]。文献[12]中作者把移动社交网络中的资料匹配问题转化为计算向量的 L_1-距离，设计了 3 个基于 Paillier 同态加密算法的细粒度私密匹配协议以此实现不同的隐私保护需求。这样的协议可以在不暴露用户私有信息的前提下，匹配用户间的个人资料数据，从而为用户间的协作建立提供方便。文献[13]中提出了 2 个基于多项式秘密共享以及隐秘集合交集技术[14]的分布式隐私安全资料匹配协议。需要指出的是，现有的隐私安全匹配多以同态加密算法为基础，其执行效率不高。通过该隐私安全资料匹配协议，协议的发起者能够在一组用户中找到与其属性最为接近的用户。尽管 Lu 等[17]在 SPOC 框架中提出了一种不依赖于同态加密算法的隐私内积计算协议来实现匹配，但是该协议较为复杂且存在明显的隐私漏洞，即当用户属性数目不大时，协议发起者很容易就能穷举出其他参与者的资料信息。

因此，设计一个能够满足上述匹配类型且效率较高的匹配协议显得尤为重要。与前述工作类似，本文把上述匹配问题转化为隐私内积计算问题。本文针对不同隐私要求设计了 3 个不依赖同态加密的隐私内积计算协议并对协议的正确性、安全性进行了理论分析。除此之外，本文还将文中所提出的 3 个协议的计算开销、通信开销与现有工作进行了理论上的比较。同时，仿真结果表明文中所提协议能够高效地完成隐私安全匹配。

2 预备知识

2.1 系统模型

这里考虑一个典型的移动自组织网络，如图 1 所示，该网络主要由一个可信机构（TA）以及众多用户组成。这里的可信机构负责系统的管理，例如系统用户的注册管理、认证等。记该系统中用户的全集为 $P=\{P_1,P_2,P_3,\cdots\}$。系统中的每个用户 P_i 都拥有 n 个属性，在不同的情况下属性有不同的含义，例如兴趣、病症、可用应用等。这里假定用户相应的个人属性档案均经过规范化排序处理，并用一个 d 维向量表示，记为 $\boldsymbol{S_i}=(S_{i,1},S_{i,2},S_{i,3},\cdots,S_{i,d})$。这里 $S_{i,k}=0$ 或者 1，$S_{i,k}$ 的值表示用户 P_i 在第 k 个属性的有无或者能力。

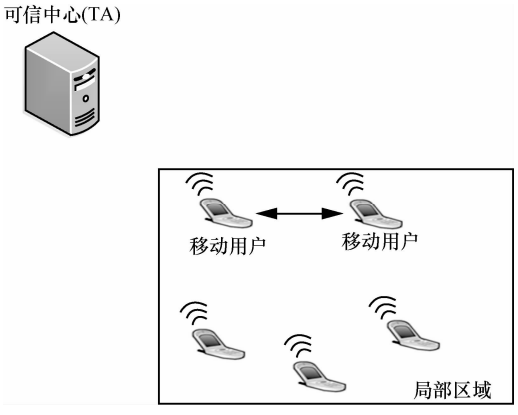


图 1 系统模型

类似于以前的研究工作[7,12,13]，本文假定该系统中协议的参与者都是 HBC(honest-but-curious)的，

即他们能够忠实地执行既定协议，但是会根据自身所掌握的信息来推测其他用户的属性档案信息或者某些特别属性，并且该系统中的用户不会恶意地联合其他用户进行合谋攻击。同时，本文中只关注用户间隐私安全的选择性匹配，因此，诸多外部攻击诸如 DoS 攻击、重放攻击以及侧信道攻击等本文未加以讨论。

另外，假定该系统中的用户还拥有下述 2 种属性。

1) 移动性。假设用户可以在一定区域内运动，只有用户在通信允许范围内时才能通过蓝牙、Wi-Fi 等进行通信。

2) 活跃性。假设该系统中的用户在有用户接近时，倾向于与他人建立通信来分享经验或者寻求帮助。

2.2 问题描述及相关定义

本文假设协作者匹配主要有下述基本的隐私保护需求：用户间在进行匹配时，用户的个人属性档案信息应受到保护，即匹配未成功的用户仅知道匹配未成功，但不了解对方的属性档案信息。这里考虑匹配发起用户 P_i 以及匹配应答用户 P_j，记 P_i 和 P_j 的个人属性档案分别为 $\boldsymbol{S}_i = (S_{i,1},S_{i,2},S_{i,3},\cdots,S_{i,d})$ 和 $\boldsymbol{S}_j = (S_{j,1},S_{j,2},S_{j,3},\cdots, S_{j,d})$。令函数 $f(\cdot,\cdot)$为一匹配度量函数。

基于不同匹配类型的不同隐私性考虑，并结合现有工作对隐私匹配问题的定义，这里给出机会网络中 3 种最为常见的隐私匹配需求。

定义 1　Level-1 隐私匹配。匹配过程中用户的个人属性信息应受到保护。协议结束时，用户的属性信息只被匹配成功的用户了解，同时匹配发起者 P_i 能够了解到其请求信息与其他用户属性档案的共同属性。

定义 2[12]　Level-2 隐私匹配。匹配过程中用户的个人属性档案信息应受到保护。协议结束时，用户的属性信息只被匹配成功的用户了解，同时，匹配发起者 P_i 能够了解到其请求信息与其他用户属性档案的整体相似度，即共同属性的个数。

定义 3[12]　Level-3 隐私匹配。匹配过程中用户的个人属性档案信息应受到保护。协议结束时，用户的属性信息只被匹配成功的用户了解，同时，匹配发起者 P_i 得不到其请求信息与其他用户属性档案的整体相似度，但能够了解到其整体相似度是否大于某一阈值。

为解决上述 3 种程度不同的隐私匹配问题，本文将给出 3 个对应的匹配协议。

3 Level-1 隐私匹配协议

3.1 Level-1 隐私匹配协议设计

在 Level-1 隐私匹配协议中，本文定义该协议所用到的匹配度量函数如下。

定义 4　Level-1 匹配度量函数。为使匹配发起者 P_i 能够了解到其请求信息与其他用户属性档案的共同属性，现引入如下带权内积

$$f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)=\sum_{k=1}^{d} o_k S_{i,k} S_{j,k}$$

其中，$\boldsymbol{O}=(o_1,o_2,o_3,\cdots,o_d)\in \mathbf{Z}^d$，并且满足 $o_1>0$，对于不小于 2 的 k 有 $o_k > \sum_{i=1}^{k-1} o_i$。

根据定义 4 很容易得到下述引理。

引理 1　对于 $1\leqslant n<d-1$，若 $o_n \leqslant f_1(\boldsymbol{S}_i, \boldsymbol{S}_j)< o_{n+1}$，则必有 $S_{i,n}= S_{j,n}=1$;并且当 $f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)\geqslant o_d$ 时，必有 $S_{i,d}= S_{j,d}=1$。

证明　当 $f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)\geqslant o_d$ 时，必有 $S_{i,d}= S_{j,d}=1$ 是明显的。对于 $1\leqslant n<d-1$，因为此时 $f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)< o_{n+1}$，对于任意 $k\geqslant n$, $S_{i,k}$ 和 $S_{j,k}$ 是明显不同时为 1 的。此时 $f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)=\sum_{k=1}^{d} o_k S_{i,k} S_{j,k}=\sum_{k=1}^{n} o_k S_{i,k} S_{j,k}$。若假设 $S_{i,n}= S_{j,n}=1$ 不成立，则有 $f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)=\sum_{k=1}^{n-1} o_k S_{i,k} S_{j,k} \leqslant \sum_{k=1}^{n-1} o_k < o_n$。与已知矛盾。因此 $S_{i,n}= S_{j,n}=1$。证毕。

由引理 1 的证明可知，当匹配发起者 P_i 得到 Level-1 匹配度量函数 $f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)$ 的值后，P_i 首先找到一个 n 使 $o_n \leqslant f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)< o_{n+1}$，以此确定用户 P_i 和用户 P_j 所共有的第 n 个属性；然后令 $f_1(\boldsymbol{S}_i,\boldsymbol{S}_j) = f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)- o_n$，重复上述步骤，直至 $f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)=0$。该过程如算法 1 所示。

算法 1

输入　Level-1 匹配度量函数值 $f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)$、一个空集合 S 以及权值向量 $\boldsymbol{O}=(o_1,o_2,o_3,\cdots,o_d)$。

输出　发起者 P_i 的请求信息 $\boldsymbol{S_i}$ 与匹配应答用户 P_j 属性档案 $\boldsymbol{S_j}$ 的共同属性集合 S。

步骤 1　判断 $f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)\geqslant o_d$ 是否成立。若是，则令 $S=S\cup\{d\}$ 且 $f_1(\boldsymbol{S}_i,\boldsymbol{S}_j) = f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)- o_d$，执行步骤 2。反之，直接执行步骤 2。

步骤 2　若 $f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)=0$，输出集合 S。反之，执行步骤 3。

步骤 3　寻找一个 n 使 $o_n \leqslant f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)< o_{n+1}$，并令

S=S∪{d}且 $f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)=f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)-o_n$。返回步骤 2。

由上述分析可以知道，实现 Level-1 隐私匹配的关键在于权值向量的选取。为简单起见可以令权值向量 $\boldsymbol{O}$ =(1,2,2^2,…,2^{d-1})。下面将给出 Level-1 匹配度量函数的隐私计算协议，协议的具体步骤如下。

1) 用户 P_i 随机选择一个 l（l >d+2）位的大整数 m 以及一个至少(d+2l+1)位的大整数 M，满足 M>(2o_d+1)m^2。

2) 用户 P_i 随机选取 d 个随机数(a_1,a_2,a_3,…,a_d)，使 $\sum_{i=1}^{d}a_i<\frac{m}{2}$；同时随机选取一组 λ 位的整数(r_1,r_2,r_3,…,r_d)。

3) 对于 S_i 的每一个元素 $S_{i,k}$ (1≤k≤d),计算 $e_k=o_kS_{i,k}m+r_kM+a_k$ 并将向量 e = (e_1,e_2,e_3,…,e_d) 以及 m 发送给用户 P_j。

4) 用户 P_j 在收到 P_i 的信息后，首先生成 d 个随机数(b_1,b_2,b_3,…,b_d)，满足对任意的1≤k≤d，有 $b_k<\frac{m}{2^{d+1}+1}$；然后计算 F=$\sum_{k=1}^{d}e_k(mS_{j,k}+b_k)$。$P_j$ 发送 F 的值给用户 P_i。

5) 用户 P_i 收到 P_j 的消息后计算 $\left\lfloor\frac{F \bmod M}{m^2}\right\rfloor$，则该值即为 $f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)$ 的值。

6) 用户根据 $f_1(\boldsymbol{S}_i,\boldsymbol{S}_j)$ 的值，利用算法 1 计算发起者 P_i 的请求信息 $\boldsymbol{S}_i$ 与匹配应答用户 P_j 属性档案 $\boldsymbol{S}_j$ 的共同属性集合 S。

3.2 Level-1 隐私匹配协议分析

上述分析可以看到，Level-1 隐私匹配协议不依赖于任何同态加密算法。下面将针对该协议的正确性、安全性以及计算和通信开销进行分析。

首先需要指出的是，文中所提出的 Level-1 隐私匹配协议是正确的。步骤 1)~步骤 3)仅涉及对原始数据的混淆，这里只需证明步骤 4)及步骤 5)的正确性即可。由步骤 4)可知，

$$\begin{aligned}F&=\sum_{k=1}^{d}e_k(mS_{j,k}+b_k)\\&=\sum_{k=1}^{d}(o_kS_{i,k}m+r_kM+a_k)(mS_{j,k}+b_k)\\&=\sum_{k=1}^{d}(o_kS_{i,k}S_{j,k}m^2+r_kmMS_{j,k}+a_kmS_{j,k})+\\&\quad\sum_{k=1}^{d}(b_ko_kS_{i,k}m+b_kr_kM+a_kb_k)\\&=\sum_{k=1}^{d}o_kS_{i,k}S_{j,k}m^2+\sum_{k=1}^{d}(r_kmS_{j,k}+r_kb_k)M+\\&\quad\sum_{k=1}^{d}(a_kmS_{j,k}+a_kb_k+b_ko_kS_{i,k}m)\end{aligned}$$

由于 $\sum_{i=1}^{d}a_i<\frac{m}{2}$，明显有

$$\sum_{k=1}^{d}a_kmS_{j,k}<\frac{m^2}{2}$$

又由于对任意的1≤k≤d，有 $b_k<\frac{m}{2^{d+1}+1}$，所以 $\sum_{k=1}^{d}(a_kb_k+b_ko_kS_{i,k}m)<\frac{m}{2^{d+1}+1}\sum_{k=1}^{d}a_k+\frac{m^2}{2^{d+1}+1}\sum_{k=1}^{d}o_k<\frac{m}{2^{d+1}+1}\frac{m}{2}+\frac{m^2}{2^{d+1}+1}(2^d-1)<\frac{m^2}{2}$，而 $\sum_{k=1}^{d}o_kS_{i,k}S_{j,k}m^2\leqslant m^2\sum_{k=1}^{d}o_k<2m^2o_d$，所以有 $\sum_{k=1}^{d}o_kS_{i,k}S_{j,k}m^2+\sum_{k=1}^{d}(a_kmS_{j,k}+a_kb_k+b_ko_kS_{i,k}m)<(2o_d+1)m^2<M$。因此，$F \bmod M=\sum_{k=1}^{d}o_kS_{i,k}S_{j,k}m^2+\sum_{k=1}^{d}(a_kmS_{j,k}+a_kb_k+b_ko_kS_{i,k}m) \bmod M$。明显地，

$$\begin{aligned}&\left\lfloor\frac{F \bmod M}{m^2}\right\rfloor\\&=\left\lfloor\frac{\sum_{k=1}^{d}o_kS_{i,k}S_{j,k}m^2+\sum_{k=1}^{d}(a_kmS_{j,k}+a_kb_k+b_ko_kS_{i,k}m)}{m^2}\right\rfloor\\&=\sum_{k=1}^{d}o_kS_{i,k}S_{j,k}=f_1(S_i,S_j)\end{aligned}$$

步骤 6)的正确性已证明。综上可知，文中所提出的 Level-1 隐私匹配协议是正确的。

其次，文中提出的 Level-1 隐私匹配协议是安全的，也就是说该协议能够保护参与者的请求信息和属性信息不被泄露。这里参与用户 P_i 仅能以很小的概率猜测到应答用户 P_j 的全部属性信息，即 Level-1 隐私匹配协议是安全的。步骤 3)中用户 P_i 发送向量 $\boldsymbol{e}$ = (e_1,e_2,e_3,…,e_d)以及 m 发送给用户 P_j。借鉴文献[17]中的结论，对于每一个 e_k(1≤k≤d)，由于用户 P_j 不知道随机数 r_kM 和 a_k，因此用户 P_j 无法区分 e_k 是 $o_km+r_kM+a_k$ 或者 r_kM+a_k 中具体哪一种形式，也就是说 P_j 无法准确得知 $S_{i,k}$ 的值。同时，又由于随机数对(r_kM,a_k)每次协议仅使用一次，因此 P_j 也无法通过多次协议之间的 e_k 确定 $S_{i,k}$ 的值。另一方面，对于用户

P_j，P_j 计算 $\sum_{k=1}^{d} e_k(mS_{j,k}+b_k)$ 并发送给用户。注意到，若没有混淆量 b_k，当用户属性数目 d 很小时，用户 P_i 只需要穷举 2^d 次便可得到用户 P_j 的属性信息。因此加入混淆量 b_k 是必要的。由于用户 P_i 不知道 $S_{j,k}$ 和 b_k 的值，此时用户 P_i 能够得到用户 P_j 的个人属性档案 S_j 需要进行 $\left(\frac{2m}{2^{d+1}+1}\right)^d$ 次穷举，因此 P_i 能够得到用户 P_j 的个人属性档案 S_j 的概率小于 $\left(\frac{2^d+1}{m}\right)^d$，当用户属性数目 d 不大时可以通过限制选取合适的 m 来加大穷举的难度。因此，当该穷举足够难时，用户 P_i 无法准确得到用户 P_j 的个人属性档案 S_j。

下面将分析 Level-1 隐私匹配协议的计算开销以及通信开销。由于随机数可以在使用之前生成并且大多数同态加密算法也涉及生成大随机数，因此本节中将不考虑生成随机数的时间开销。同时，加运算相较乘运算的开销要小得多，因此 Level-1 隐私匹配协议的计算开销可近似为大整数乘运算的次数。可以看到用户 P_i 需要进行最多 d 次 l 位乘运算以及 d 次(d+2l+1)位乘运算以及 1 次(d+2l+1)位模运算和 1 次 2l 位除运算，而用户 P_j 需要进行至多 d 次(d+2l+λ+1)位乘运算。在该协议中，用户 P_i 需要发送(d (d+2l+λ+1)+ l)位数据给用户 P_j，与此同时用户 P_j 需要发送(d+3l+λ+1)位数据给用户 P_i。

4　Level-2 隐私匹配协议

4.1　Level-2 隐私匹配协议设计

Level-2 隐私匹配协议执行结束时，匹配发起者 P_i 仅需要了解到其请求信息与其他用户属性档案的整体相似度，即共同属性的个数。为实现这一目标，本文定义该协议所用到的匹配度量函数如下。

定义 5　Level-2 匹配度量函数。为使匹配发起者 P_i 能够了解到其请求信息与其他用户属性档案的共同属性数目，这里只要引入下述标准内积函数

$$f_2(\boldsymbol{S}_i,\boldsymbol{S}_j)=\sum_{k=1}^{d} S_{i,k}S_{j,k}$$

这里只有当 $S_{i,k}=S_{j,k}=1$，即请求信息与其他用户属性档案同时拥有第 n 个属性时，$S_{i,k}S_{j,k}=1$。

下面将给出 Level-2 匹配度量函数的隐私计算协议。协议的具体步骤如下。

1) 用户 P_i 随机选择一个 l（l>$\lfloor \mathrm{lb}(2d+1) \rfloor$ +2）位的大整数 m 以及一个至少 $(\lfloor \mathrm{lb}(d+1) \rfloor+2l+1)$ 位的大整数 M，满足 M>(d+1)m^2。

2) 用户 P_i 随机选取 d 个随机数(a_1,a_2,a_3,⋯,a_d)，使 $\sum_{i=1}^{d} a_i < \frac{m}{2}$；同时随机选取一组 λ 位的整数(r_1,r_2,r_3,⋯,r_d)。

3) 对于 S_i 的每一个元素 $S_{i,k}$ ($1 \leqslant k \leqslant d$)，计算 $e_k = S_{i,k}m + r_k M + a_k$，并将向量 $\boldsymbol{e}$=(e_1,e_2,e_3,⋯,e_d) 以及 m 发送给用户 P_j。

4) 用户 P_j 在收到 P_i 的信息后，首先生成 d 个随机数(b_1,b_2,b_3,⋯,b_d)，满足对任意的$1 \leqslant k \leqslant d$，有 $b_k < \frac{m}{2d+1}$；然后计算 F=$\sum_{k=1}^{d} e_k(mS_{j,k}+b_k)$。$P_j$ 发送 F 的值给用户 P_i。

5) 用户 P_i 收到 P_j 的消息后计算 $\left\lfloor \frac{F \bmod M}{m^2} \right\rfloor$，则该值即为 $f_2(\boldsymbol{S}_i,\boldsymbol{S}_j)$ 的值。

4.2　Level-2 隐私匹配协议分析

Level-2 隐私匹配协议的正确性以及隐私保护性的证明与 3.2 节中的证明类似，这里不做赘述。注意到在 Level-2 隐私匹配协议的第 4)步中，这里同样引入了一个混淆量 b_k，此时由于 P_i 不知道 $S_{j,k}$ 和 b_k 的值，因此用户 P_i 能够得到用户 P_j 的个人属性档案 S_j 的概率为 $\left(\frac{2d+1}{2m}\right)^d$。下面主要分析该协议的计算开销与通信开销。可以看到用户 P_i 需要进行最多 d 次 $(\lfloor \mathrm{lb}(d+1) \rfloor+2l+1)$ 位乘法运算以及 1 次 $(\lfloor \mathrm{lb}d+1) \rfloor+2l+1)$ 位模运算和 1 次 2l 位除法运算，而用户 P_j 需要进行至多 d 次 $(\lfloor \mathrm{lb}(d+1) \rfloor+2l+\lambda+1)$ 位乘运算。在该协议中，用户 P_i 需要发送 d ($\lfloor \mathrm{lb}(d+1) \rfloor$+2$l$+$\lambda$+1)+$l$)位数据给用户 P_j，与此同时用户 P_j 需要发送($\lfloor \mathrm{lb}(d+1) \rfloor$+3$l$+$\lambda$+1)位数据给用户 P_i。

5　Level-3 隐私匹配协议

5.1　Level-3 隐私匹配协议设计

Level-3 隐私匹配协议要求协议结束时，用户的属性信息只被匹配成功的用户了解，同时，匹配发起者 P_i 得不到其请求信息与其他用户属性档案的整体相似度，但能够了解到其整体相似度是否大于某一阈值 τ。该协议中仍然使用定义 5 中的匹配度量函数，其具体步骤如下。

1) 用户 P_i 随机选择一个 l 位的大整数 m、一个 κ 位的大整数 δ 以及一个至少 $(\lfloor \mathrm{lb}(d+1) \rfloor + 2l + \kappa + 1)$ 位的大整数 M，满足 $M > \delta\,(d+2)m^2$。

2) 用户 P_i 随机选取 d 个随机数 $(a_1, a_2, a_3, \cdots, a_d)$，使 $\sum_{i=1}^{d} a_i < \frac{m}{2}$；同时随机选取一组 λ 位的整数 $(r_1, r_2, r_3, \cdots, r_d)$。

3) 对于 $\boldsymbol{S}_i$ 的每一个元素 $S_{i,k}$ $(1 \leqslant k \leqslant d)$，计算 $e_k = S_{i,k} m + r_k M + a_k$ 并将向量 $e=(e_1, e_2, e_3, \cdots, e_d)$、$\delta$、$m$ 以及阈值 τ 发送给用户 P_j。

4) 用户 P_j 在收到 P_i 的信息后，计算 $F=\sum_{k=1}^{d} e_k m S_{j,k}$，然后用户 P_j 生成 3 个随机数 δ_1、δ_2、δ_3，使 $\delta \geqslant \delta_1 > \delta_2 > \delta_3 \geqslant 0$ 并且满足 $\delta_3 \bmod \tau = 0$ 以及 $\frac{\delta_2 - \delta_3}{\delta_1} < \frac{1}{2}$。$P_j$ 计算 $\theta_1 = \delta_1 F + \delta_2 m^2$ 和 $\theta_2 = (\delta_1 \tau + \delta_3) m^2$。$P_j$ 发送 θ_1 以及 θ_2 的值给用户 P_i。

5) 用户 P_i 收到 P_j 的消息后计算 $\theta_3 = \theta_1 \bmod M$，若 $\theta_3 > \theta_2$，则有 $f_2(\boldsymbol{S}_i, \boldsymbol{S}_j) \geqslant \tau$；反之 $f_2(\boldsymbol{S}_i, \boldsymbol{S}_j) < \tau$。

5.2 Level-3 隐私匹配协议分析

上述分析可以看到，Level-3 隐私匹配协议不同于其他 2 个协议。本节将针对该协议的正确性、安全性以及计算和通信开销进行分析。

首先需要指出的是 Level-3 隐私匹配协议是正确的。由步骤 4)可知，

$$F=\sum_{k=1}^{d} e_k m S_{j,k} = \sum_{k=1}^{d} m S_{j,k} (S_{i,k} m + r_k M + a_k)$$

$$= \sum_{k=1}^{d} (S_{j,k} S_{i,k} m^2 + S_{j,k} r_k m M + S_{j,k} a_k m)$$

因此，

$$\theta_1 = \delta_1 F + \delta_2 m^2$$

$$= (\delta_1 f_2(\boldsymbol{S}_i, \boldsymbol{S}_j) + \delta_2) m^2 + m M \delta_1 \sum_{k=1}^{d} S_{j,k} r_k + m \delta_1 \sum_{k=1}^{d} S_{j,k} a_k$$

所以有

$$\theta_3 = \theta_1 \bmod M = (\delta_1 f_2(\boldsymbol{S}_i, \boldsymbol{S}_j) + \delta_2) m^2 + m \delta_1 \sum_{k=1}^{d} S_{j,k} a_k$$

由于 $\sum_{i=1}^{d} a_i < \frac{m}{2}$，因此

$$\theta_3 < (\delta_1 f_2(\boldsymbol{S}_i, \boldsymbol{S}_j) + \delta_2 + \frac{1}{2}) m^2$$

而 $\theta_2 = (\delta_1 \tau + \delta_3) m^2$，此时

$$\frac{\theta_3 - \theta_2}{\delta_1 m^2} < f_2(\boldsymbol{S}_i, \boldsymbol{S}_j) - \tau + \frac{\delta_2 - \delta_3}{\delta_1} + \frac{1}{2}$$

因为 $f_2(\boldsymbol{S}_i, \boldsymbol{S}_j)$ 为自然数，而 $\frac{\delta_2 - \delta_3}{\delta_1} < \frac{1}{2}$，所以当 $\theta_3 > \theta_2$ 时，$f_2(\boldsymbol{S}_i, \boldsymbol{S}_j) - \tau + \frac{\delta_2 - \delta_3}{\delta_1} + \frac{1}{2} > 0$，即 $f_2(\boldsymbol{S}_i, \boldsymbol{S}_j) \geqslant \tau$；反之，$f_2(\boldsymbol{S}_i, \boldsymbol{S}_j) - \tau \leqslant -1$，故此时 $f_2(\boldsymbol{S}_i, \boldsymbol{S}_j) < \tau$。至此 Level-3 隐私匹配协议的正确性得证。

其次，需要证明文中提出的 Level-3 隐私匹配协议能够保护参与者的请求信息和属性信息不被泄露，并且用户 P_i 仅能以很小的概率猜测到其请求信心与应答用户 P_j 资料属性的整体相似程度。这里 P_j 无法准确得知 $S_{i,k}$ 的值类似于 3.2 节中的证明。另一方面，在步骤 4)中，用户 P_j 将 $\theta_1 = \delta_1 F + \delta_2 m^2$ 以及 $\theta_2 = (\delta_1 \tau + \delta_3) m^2$ 的值发送给用户 P_i，由于用户 P_i 不知道 δ_1、δ_2 和 δ_3 的值，因此，此时用户 P_i 能够得到用户 P_j 的个人属性档案 S_j 的概率为 $\frac{1}{2^{d-1} \delta^3}$。

对于 Level-3 隐私匹配协议的计算开销和通信开销，可以看到用户 P_i 需要进行最多 d 次 $(\lfloor \mathrm{lb}(d+1) \rfloor + 2l + \kappa + 1)$ 位乘法运算以及 1 次 $(\lfloor \mathrm{lb}(d+1) \rfloor + 2l + \kappa + 1)$ 位模运算，而用户 P_j 需要进行至多 $(d+1)$ 次 $(\lfloor \mathrm{lb}(d+1) \rfloor + 2l + \kappa + 1)$ 位乘法运算以及 3 次 $2l$ 位乘法运算。在该协议中，用户 P_i 需要发送 $(d(\lfloor \mathrm{lb}(d+1) \rfloor + 2l + \kappa + 1) + l + \lfloor \mathrm{lb}(d+1) \rfloor + \kappa)$ 位数据给用户 P_j，与此同时用户 P_j 需要发送 $(2\lfloor \mathrm{lb}(d+1) \rfloor + 5l + 2\kappa + 1)$ 位数据给用户 P_i。

6 性能评估

本文所设计的隐私安全资料匹配协议不依赖于同态加密算法，因此其计算效率相对于以同态加密算法为基础的匹配协议具有较高的计算效率。在 6.1 节中，本文选择 Lu 等[17]的 PPSPC 协议以及 Zhang 等[12]提出的协议 2 进行理论上的性能比较。在 6.2 节中，通过仿真实验对比了文中所提出的 Level-2 隐私匹配协议与 Zhang 等的协议 2 的真实计算开销与通信开销。

6.1 性能对比

首先，Lu 等在 SPOC 框架中提出了一种不依赖于同态加密算法的隐私内积计算协议（PPSPC 协议）来实现匹配，但是该协议较为复杂且存在明显的隐私漏洞，即当用户属性数目不大时，协议发起者很容易就能穷举出其他参与者的资料信息。本文所设计的 3 个协议很好地解决了这个安全问题。在

文献[12]的协议 2、PPSPC 协议以及文中 3 个协议中，协议发起者能够穷举出其他参与者资料信息的概率如表 1 所示。这里假设 l=512，κ=128。由表 1 可以看到，文中设计的 3 个隐私安全匹配协议在用户属性数目很小时也有很好的安全性，并且 Level-3 隐私匹配协议还可以通过限制 κ 或者采取 Level-1 以及 Level-2 隐私匹配协议中同样的混淆方式来增加发起者的穷举难度。

为方便表述，这里仅考虑乘法（或者除法）以及幂运算的时间开销并定义 *mul*、*exp* 分别为乘法以及幂运算的时间开销，并且令协议 2 中明文空间也为 l 位。表 2 归纳了文中所提出的 3 个隐私安全匹配协议与 PPSPC 协议以及协议 2 的计算与通信开销。由表 2 可以看到，5 个协议的通信开销大致相同，但是由于幂运算的实际运算开销要远大于乘法运算（约 240 倍）[19]，因此理论上来说基于 Paillier 同态加密算法的匹配协议 2 的计算量要比本文中所提出的协议的计算量大 240 倍。

6.2　仿真模拟

本文通过仿真实验来对 Level-2 隐私匹配协议与 Zhang 等[12]提出的协议 2 的计算执行时间以及通信量进行比较。本节中用到的移动设备为 Lenovo A378T，其RAM为512 MB，CPU为MT6572 1.33 GHz，Java 版本为 4.2。这里未考虑设备间的通信，协议过程均在同一设备上进行。

对于 Level-2 隐私匹配协议，本节中选取 2 组参数进行测试,其中第一组参数 Para1 为 l=512，λ=256，a_i 和 b_i 均为 128 bit 的整数；第二组参数 Para2 为 l=1 024，λ=512，a_i 和 b_i 均为 256 bit 的整数。与此相对应，协议 2[12]中同样选取两组参数，其中第一组参数 Para3 中 N 为 512 bit 整数，g=2；第二组参数 Para4 中 N 为 1 024 bit 整数，g=2。在模拟时，发起者与参与者的属性资料均随机生成。这里针对不同的参数，分别执行两协议 200 次并统计两协议中发起者的加密耗时、参与者计算耗时、发起者解密耗时以及总的通信位数，其结果如图 2 所示。

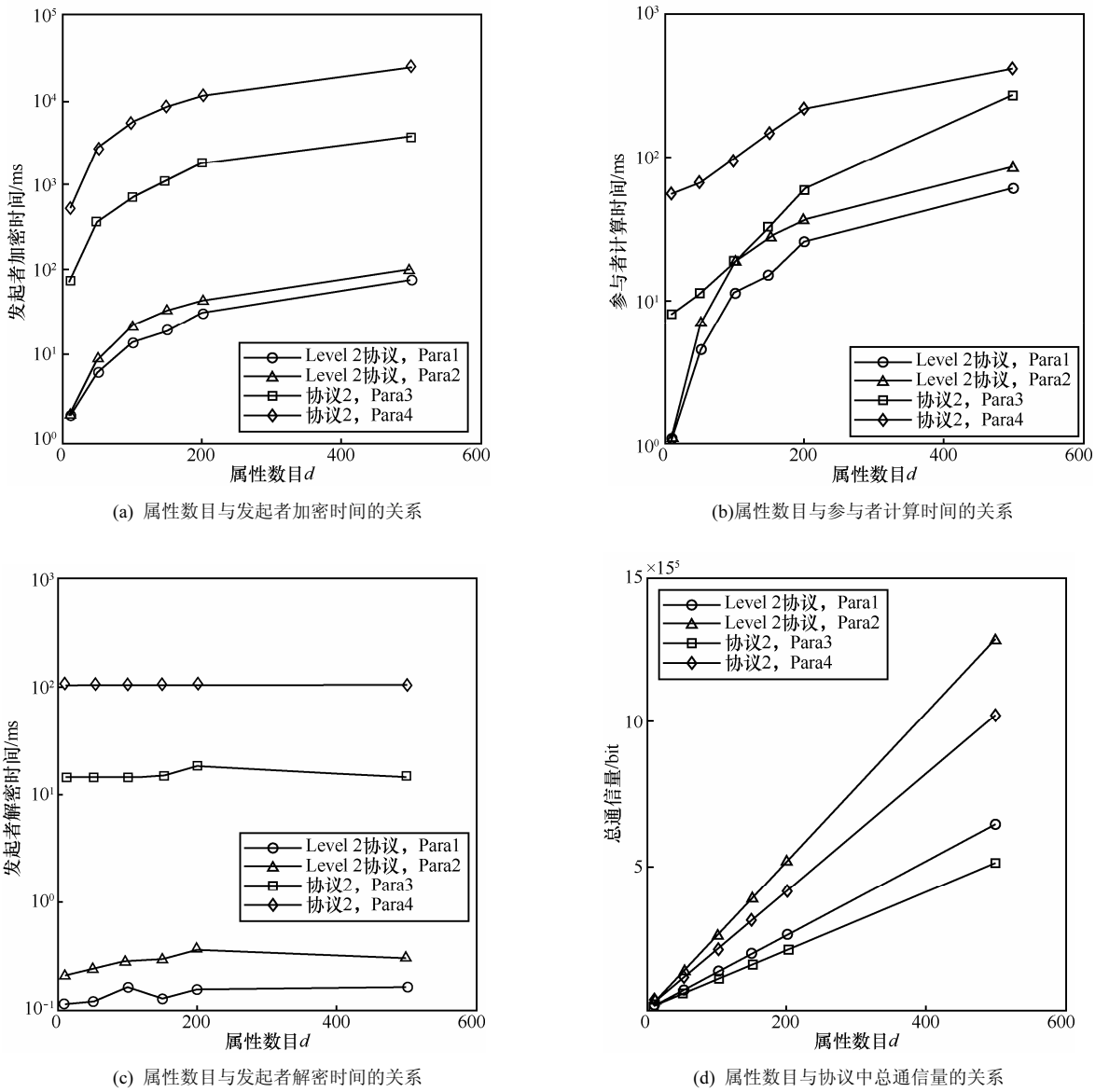
由图 2(a)可以看到，当两协议的密文位数大致相同时（两协议相应地选择 Para1 和 Para3，Para2 和 Para4），Level-2 隐私匹配协议中发起者加密数据的平均耗时远小于协议 2 的加密耗时（其中 512 bit 密钥下约为 $\frac{1}{50}$，1 024 bit 密钥下约为 $\frac{1}{240}$），并且当采用的加密数据位数变大时，Level-2 隐私匹配协议中发起者加密数据的平均耗时变化较小(约 1.5 倍)，而协议 2 的加密耗时约增加了 4 倍。由图 2(b)可知，当两协议的密文位数大致相同时（两协议相应地选择 Para1 和 Para3，Para2 和 Para4），协议 2 的参与者计算耗时约为 Level-2 隐私匹配协议耗时的 2.5 倍。由图 2(c)可以看到，2 个协议在解密阶段的效率都很高。对于发起者的解密时间，随着密文位数的增加，Level-2 隐私匹配协议的解密耗时增加了 2 倍，但其操作耗时仍小于 1 ms。对于采取参

表 1　　**发起者能够得到其他参与者资料信息的概率**

协议名称	属性数目暴露概率			
	d=10	d=20	d=50	d=100
协议 2[12]	7.28×10^{-158}	7.11×10^{-161}	6.62×10^{-170}	5.88×10^{-185}
PPSPC 协议	9.77×10^{-4}	9.54×10^{-7}	8.88×10^{-16}	7.89×10^{-31}
Level1 隐私匹配协议	$<1.00\times10^{-300}$			
Level2 隐私匹配协议	$<1.00\times10^{-300}$			
Level3 隐私匹配协议	4.96×10^{-119}	4.84×10^{-122}	4.50×10^{-131}	4.00×10^{-146}

表 2　　**开销对比**

协议名称	隐私保护级别	计算开销	通信开销
Level-1 隐私匹配协议	Level-1	$(3d+1)mul$	$(d+1)(d+2l+\lambda+1)+2l$
Level-2 隐私匹配协议	Level-2	$(2d+1)mul$	$(d+1)(\lfloor \mathrm{lb}\,(d+1)\rfloor+2l+\lambda+1)+2l$
Level-3 隐私匹配协议	Level-3	$(2d+4)mul$	$(d+1)(\lfloor \mathrm{lb}\,(d+1)\rfloor+2l+\lambda+\kappa+1)+4l+2\kappa+2\lfloor \mathrm{lb}\,(d+1)\rfloor$
协议 2[12]	Level-2	$(2d+4)exp+(2d+1)mul$	$\approx(2d+3)l$
PPSPC[15]	Level-2	$(2d+1)mul$	$(d+1)(\lfloor \mathrm{lb}\,(d+1)\rfloor+2l+\lambda+1)+2l$



(a) 属性数目与发起者加密时间的关系

(b)属性数目与参与者计算时间的关系

(c) 属性数目与发起者解密时间的关系

(d) 属性数目与协议中总通信量的关系

图 2　仿真结果对比

数 Para4 的协议 2 来讲，其解密耗时约为 Level-2 隐私匹配协议的 50 倍，其操作耗时约为 100 ms。图 2(d)描述了两协议在不同参数下总的通信开销，可以看到当两协议的密文位数大致相同时，Level-2 隐私匹配协议相对于协议 2 需要多发送大约 25%的数据。这是由于 Level-2 隐私匹配协议在上述 2 组参数下产生的密文位数较大隐私保护程度较强所致，这里本文微调选取的参数便可以消除数据量上的差异。

7　结束语

机会网络中的协作者匹配是建立机会网络的基础。考虑到用户资料属性的私密性，如何在不泄露用户资料属性的前提下完成用户间的匹配是资料匹配需要首先考虑的问题。本文针对不同隐私要求设计了 3 个不依赖同态加密的高效隐私内积计算协议。可以证明，文中所提出的协议是隐私安全并且正确的。本文还对所提出的 3 个协议的计算开销与通信开销与现有工作进行了理论上的比较。仿真结果表明文中所提协议能够高效地完成隐私安全匹配。未来在这个研究领域的工作，可以进一步考虑来自外部攻击，设计一整套完整的安全框架[20,21]；同时也可以把文中所提出的隐私安全匹配协议推广到更一般的情形以满足程度属性的匹配问题。

参考文献：

[1] CONTI M, KUMAR M. Opportunities in opportunistic computing[J]. Computer, 2010, 43(1): 42-50.

[2] SHARMA G, MAZUMDAR R, SHROFF N B. Delay and capacity trade-offs in mobile ad hoc networks: a global perspective[J]. IEEE/ACM Transactions on Networking (ToN), 2007, 15(5): 981-992.

[3] 陈曦, 李光松, 田有亮, 等. 机会网络中基于社会属性的按需密钥管理方案[J]. 通信学报, 2013, 34(12): 93-99.
CHEN X, LI G S, TIAN Y L, et al. On-demand key management based on social attribute for opportunistic networks[J]. Journal on Communications, 2013, 34(12): 93-99.

[4] CONTI M, GIORDANO S, MAY M, et al. From opportunistic networks to opportunistic computing[J]. Communications Magazine, IEEE, 2010, 48(9): 126-139.

[5] CONTI M, DAS S K, BISDIKIAN C, et al. Looking ahead in pervasive computing: challenges and opportunities in the era of cyber-physical convergence[J]. Pervasive and Mobile Commputing, 2012, 8(1): 2-21.

[6] 胡海洋, 李忠金, 胡华,等. 面向移动社交网络的协作式内容分发机制[J]. 计算机学报, 2013, 36(3):613-625.
HU H Y, LI Z J, HU H, et al. Cooperative contents distribution in mobile social networks[J]. Chinese Journal of Computers,2013,36(3): 613-625.

[7] LU R, LIN X, LIANG X, et al. A secure handshake scheme with symptoms-matching for mhealthcare social network[J]. Mobile Networks and Applications, 2011, 16(6): 683-694.

[8] KAPADIA A, KOTZ D, TRIANDOPOULOS N. Opportunistic sensing: security challenges for the new paradigm[A]. First International in Communication Systems and Networks and Workshops[C]. Bangalore, India, 2009.1-10.

[9] 熊永平, 孙利民, 牛建伟, 等. 机会网络[J]. 软件学报, 2009, 20(1): 124-137.
XIONG Y P, SUN L M, NIU J W, et al. Opportunistic networks[J]. Journal of Software, 2009, 20(1): 124−137.

[10] 徐佳, 孙力娟, 王汝传, 等. 机会网络中基于种子喷雾的自适应路由协议[J]. 电子学报, 2010, 38(10): 2315-2321.
XU J, SUN L J,WANG R C, et al. Adaptive seed spay routing for opportunistic networks[J]. Acta Electronic Sinica, 2010, 38(10): 2315-2321.

[11] WANG X, CHENG W, MOHAPATRA P, et al. Enabling reputation and trust in privacy-preserving mobile sensing[J]. IEEE Transactions on Mobile Computing, 2014, 13(12): 2777-2790.

[12] ZHANG R, ZHANG J, ZHANG Y, et al. Privacy-preserving profile matching for proximity-based mobile social networking[J]. IEEE Journal on Selected Areas in Communications, 2013, 31(9):656 - 668.

[13] LI M, CAO N, YU S, et al. FindU: privacy-preserving personal profile matching in mobile social networks[A]. INFOCOM[C]. Shanghai, China, 2011. 2435-2443.

[14] FREEDMAN M J, NISSIM K, PINKAS B. Efficient private matching and set intersection[A]. Cryptology-EUROCRYPT 2004[C]. Interlaken, Switzerland, 2004. 1-19.

[15] DONG C, CHEN L. A Fast Secure Dot Product Protocol with Application to Privacy Preserving Association Rule Mining[M]. Knowledge Discovery and Data Mining, Springer International Publishing, 2014.

[16] ZHU X, CHEN Z, CHI H, et al. Two-party and multi-party private matching for proximity-based mobile social networks[A]. IEEE International Conference on Communications (ICC)[C]. Sydney, Australia, 2014. 926-931.

[17] LU R, LIN X, SHEN X. SPOC: a secure and privacy-preserving opportunistic computing framework for mobile-healthcare emergency[J]. IEEE Transactions on Parallel and Distributed Systems, 2013, 24(3): 614-624.

[18] LIAO X, ULUAGAC S, BEYAH R A. S-match: verifiable privacy-preserving profile matching for mobile social services[A]. The 44th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)[C]. Atlanta,USA, 2014. 287-298.

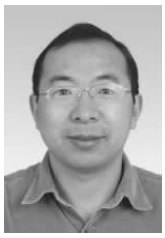
[19] HUANG K H, CHUNG Y F, LIU C H, et al. Efficient migration for mobile computing in distributed networks[J]. Computer Standards & Interfaces, 2009, 31(1): 40-47.

[20] LU R, LIN X, SHI Z, et al. PLAM: a privacy-preserving framework for local-area mobile social networks[A]. INFOCOM[C]. Toronto, Canada, 2014.763-771.

[21] KAYASTHA N, NIYATO D, WANG P, et al. Applications, architectures, and protocol design issues for mobile social networks: a survey[J]. Proceedings of the IEEE, 2011, 99(12):2130-2158.

作者简介：

李永凯（1988-），男，山东临沂人，武汉大学博士生，主要研究方向为信息安全、隐私保护等。



刘树波（1970-），男，蒙古族，黑龙江齐齐哈尔人，博士，武汉大学教授、博士生导师，主要研究方向为信息安全、隐私保护、嵌入式系统等。

杨召唤（1991-），男，河南商丘人，武汉大学硕士生，主要研究方向为隐私保护。

刘梦君（1988-），男，湖北黄冈人，武汉大学博士生，主要研究方向为信息安全、隐私保护等。